

The Direction of RecA Protein Assembly onto Single Strand DNA Is the Same as the Direction of Strand Assimilation during Strand Exchange*

(Received for publication, May 1, 1985)

James C. Register, III and Jack Griffith†

From the Lineberger Cancer Research Center and the Departments of Microbiology and Immunology, and Biochemistry, University of North Carolina Medical School, Chapel Hill, North Carolina 27514

The RecA protein of *Escherichia coli* optimally promotes DNA strand exchange reactions in the presence of the single strand DNA-binding protein of *E. coli* (SSB protein). Under these conditions, assembly of RecA protein onto single-stranded DNA (ssDNA) occurs in three steps. First, the ssDNA is rapidly covered by SSB protein. The binding of RecA protein is then initiated by nucleation of a short tract of RecA protein onto the ssDNA. Finally, cooperative polymerization of additional RecA protein accompanied by displacement of SSB protein results in a ssDNA-RecA protein filament (Griffith, J. D., Harris, L. D., and Register, J. C. (1984) *Cold Spring Harbor Symp. Quant. Biol.* 49, 553-559). We report here that RecA protein assembly onto circular ssDNA yields RecA protein-covered circles in which greater than 85% are completely covered by RecA protein with no remaining SSB protein-covered segments (as detected by electron microscopy). However, when linear ssDNA is used, 90% of the filaments contain a short segment at one end complexed with SSB protein. This suggests that RecA protein assembly is unidirectional. Visualization of the assembly of RecA protein onto either long ssDNA tails (containing either 5' or 3' termini) or ssDNA gaps generated in double strand DNA allowed us to determine that the RecA protein polymerizes in the 5' to 3' direction on ssDNA and preferentially nucleates at ssDNA-double strand DNA junctions containing 5' termini.

The best available *in vitro* model system for homologous recombination is the strand exchange reaction promoted by the RecA protein of *Escherichia coli* (for review, see Ref. 1). In the simplest reactions, a (+)-ssDNA¹ pairs with a dsDNA in a region of shared homology and the two molecules undergo a net exchange of DNA strands (2-4). One of the hallmarks of this exchange is that it proceeds unidirectionally, displacing the 5' end of the (+)-strand of the dsDNA partner (5-7). Despite the importance of this polarity in strand exchange

(for review, see Ref. 8), a satisfactory explanation for this phenomenon has yet to be offered.

Strand exchange occurs in three experimentally separable phases: formation of initiation complexes, synapsis, and branch migration. The cooperative binding of RecA protein to ssDNA in the presence of ATP and Mg²⁺ forms an initiation (or presynaptic) complex (9-14). This is followed by synapsis in which a stable joint forms between the ssDNA and dsDNA molecules within a region of homology (1, 15-18). Movement of this joint (branch migration) results in a net exchange of DNA strands (2-4). Branch migration proceeds unidirectionally in the 3'→5' direction with respect to the (-)-DNA strand of the dsDNA (the strand being paired with the ssDNA) (5-8).

Initiation complexes formed by RecA protein in the presence of SSB protein are 5-10-fold more active in strand exchange than those formed in the absence of SSB protein (2, 19, 20). As visualized by electron microscopy (14, 21), these complexes are smooth rigid filaments which are easily distinguished from the nucleosomal ssDNA·SSB protein complexes (22). Using electron microscopy, we have shown (21) that when RecA protein, SSB protein, Mg²⁺, and ATP are added simultaneously to a circular ssDNA template (SSB-mediated RecA protein assembly), SSB protein binds first. RecA protein binding is then initiated by the relatively rare nucleation of RecA protein on the SSB protein-covered ssDNA. A rapid, highly cooperative polymerization of RecA protein along the ssDNA then ensues, resulting in the formation of initiation complexes. We have not observed any evidence for a site-specific nucleation of RecA protein onto ssDNA and therefore postulate that such nucleation occurs at random sites. Studies with labeled proteins indicate that, under the conditions used in our studies, as RecA protein assembly proceeds, most if not all of the SSB protein is displaced.²

Recently, we noted that when linear ssDNA was used as a template for RecA protein assembly, a short segment of one end of the initiation complex almost always exhibited the nucleosomal appearance typical of ssDNA which is covered by SSB protein. This result suggested that the assembly of RecA protein was unidirectional. To test this hypothesis, we examined the assembly of RecA protein onto dsDNA templates having ssDNA tails with 3' or 5' termini. In this paper, we demonstrate, that under optimal strand exchange conditions, RecA protein binds to ssDNA in the 5'→3' direction. We also show that SSB-mediated RecA protein assembly onto ssDNA will preferentially nucleate at a ssDNA-dsDNA junction containing a 5' terminus if it is present.

* This work was supported by Grant GM 31819 from the National Institutes of Health and National Cancer Institute Grant CA 16086. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed.

¹ The abbreviations used are: ssDNA, single strand DNA; dsDNA, double strand DNA; SSB protein, single strand DNA-binding protein of *E. coli*; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.

² R. Thresher, S. Colby, G. Christiansen, and J. Griffith, manuscript in preparation.

MATERIALS AND METHODS

Proteins—RecA protein was purified from *E. coli* overproducer strain KM4104/pDR1453 as previously described (23). SSB protein was purified as described by Chase *et al.* (24) from an *E. coli* overproducer strain (gift of J. W. Chase, Department of Molecular Biology, Albert Einstein College of Medicine). *E. coli* exonuclease III was purified as described by Rogers and Weiss (25).

DNA—Single-stranded DNA was purified from M13 phage yielding M13mp7 ssDNA. Following phage disruption by chloroform and 0.5% Sarkosyl treatment, the DNA was purified by banding in CsCl density gradients. This DNA contains a hairpin structure that generates a duplex region with an *EcoRI* restriction site. The ssDNA was cut by the addition of 1 unit of *EcoRI*/μg of DNA (*EcoRI* was the gift of P. Modrich, Department of Biochemistry, Duke University) and incubation at 37 °C for either 15 min to generate a mixture of linear and circular molecules or 45 min to generate only linear molecules. M13 and pUR222 dsDNA were amplified in *E. coli* K37 or HB101, respectively, by chloramphenicol treatment. Supercoiled DNA was purified by the method of Modrich and Zabel (26). M13 dsDNA was linearized by *HincII* (Bethesda Research Laboratories). *E. coli* exonuclease III treatment of linear M13 dsDNA (40 μg/ml) was carried out in 50 mM Tris, pH 8.0, 1 mM MgCl₂, and 1 mM β-mercaptoethanol buffer using 500 units/ml exonuclease III for 20 min at 21 °C. Phage λ-exonuclease (New England Biolabs) treatment of linear M13 dsDNA (40 μg/ml) was performed in 100 mM glycine, pH 9.4, 3 mM MgCl₂, and 3 mM β-mercaptoethanol buffer using 150 units of enzyme/ml for 20 min at 37 °C. Tails generated by both exonuclease III and λ-exonuclease were approximately 1000–3000 bases long. Single nicks were generated in supertwisted pUR222 by the method of Greenfield *et al.* (27). The nicks were enlarged to gaps using exonuclease III as described above. Following all nuclease treatments, the DNA was treated with 1% sodium dodecyl sulfate, extracted with phenol:chloroform (1:1), ethanol-precipitated, and resuspended in 10 mM Tris, pH 7.5. Calf intestine phosphatase (Boehringer Mannheim) was used to dephosphorylate 5' ends of dsDNA (28).

SSB-mediated Assembly of RecA Protein onto ssDNA—3 μg/ml of ssDNA was incubated with 120 μg/ml RecA protein, 4 μg/ml creatine phosphokinase (Sigma), and 20 mM phosphocreatine (Sigma) in a buffer containing 30 mM NaCl, 3 mM ATP, 12 mM MgCl₂, and 20 mM HEPES, pH 7.5, for 5 min at 37 °C. SSB protein (9 μg/ml) was then added, and incubation was continued for various times. In experiments in which incompletely polymerized RecA protein tracts were visualized, SSB protein was added to the ssDNA at the same time as the RecA protein (to slow nucleation) at 21 °C (to slow polymerization) for 2–4 min. No differences in the overall features of the assembly pathway were found whether SSB protein was added with the RecA protein or a few minutes later. The observations reported here have been confirmed using both sequences of addition.

Electron Microscopy—DNA-protein complexes were fixed with formaldehyde and glutaraldehyde, passed over 2-ml Sepharose 4B columns, mounted onto thin carbon films, and rotary-shadowcasted with tungsten (29). Micrographs were taken on a Phillips TLG 400.

RESULTS

Assembly of RecA Protein onto Linear and Circular ssDNA Templates—SSB-mediated assembly of RecA protein onto a mixture of linear and circular M13mp7 ssDNA was examined by electron microscopy. Following incubation for 10 min at 37 °C, 25% circular and 75% linear ssDNA-protein complexes were observed (Table I and Fig. 1, A and B). 88% of the circular complexes were smooth-contoured loops with a 10-nm filament diameter identical in appearance to previously described ssDNA-RecA protein complexes (9–11). There were no gaps or breaks along the length of these filaments. 12% of the circular filaments contained one long RecA protein-covered tract and one tract that retained the nucleosomal appearance of ssDNA complexed by SSB protein (21, 22) (Table I and Fig. 1, A and B). In contrast, only 8% of the linear complexes were completely covered by RecA protein, while 90% were covered by RecA protein except for 5–20% of one end which remained covered by SSB protein. 2% of the linear complexes remained fully covered by SSB protein (Table I and Fig. 1, A and B).

These results suggest that the assembly of RecA protein

TABLE I

Appearance of filaments formed by RecA protein along various DNA templates

RecA protein was incubated with several DNA templates in the presence of Mg²⁺, ATP, and SSB protein at 37 °C for either 10 min (1 and 2) or 30 min (3–5) as described under "Materials and Methods." Samples were fixed and examined by electron microscopy, and fields of molecules were scored.

DNA template	Number of molecules of each filament type		
	Fully RecA-covered	Fully SSB-covered	Partially RecA- and SSB-covered
1. Circular ssDNA	44	0	6
2. Linear ssDNA	11	2	132
3. Exonuclease III-treated dsDNA (5' termini)	4	87	31
4. λ-Exonuclease-treated dsDNA (3' termini)	150	25	85
5. Gapped DNA	72	8	7

onto ssDNA is unidirectional. If so, random nucleation of RecA protein on linear ssDNA that was covered by SSB protein would leave a segment of the ssDNA covered by SSB protein. Subsequent RecA protein nucleations within the SSB protein-covered segment would leave some ssDNA covered by SSB protein, unless the nucleation occurred at the end of the SSB protein-covered segment. However, an identical reaction along circular ssDNA would lead to full RecA protein coverage.

RecA Protein Assembly onto Single-stranded DNA Tails with 5' Termini—To test the above prediction, RecA protein was assembled onto linear M13 dsDNA which had been treated with *E. coli* exonuclease III. This DNA contained long single-stranded tails with 5' termini. Following incubations of 10–30 min, all of the molecules observed contained a central protein-free dsDNA segment with two protein-covered tails (Fig. 1, C–E). 72% of the tails appeared fully covered by SSB protein, 3% were fully covered by RecA protein, and 24% contained one segment covered by RecA protein and one segment covered by SSB protein (Table I). In this last group of molecules, the protein-covered tracts were uniquely oriented such that the SSB protein-covered segment was always at the end of the ssDNA tail (Fig. 1, C–E).

These observations could be explained in two ways: either the RecA protein nucleated randomly onto the ssDNA tail and then polymerized in the 5' to 3' direction ending at the ssDNA-dsDNA junction, or the RecA protein nucleated at the ssDNA-dsDNA junction and then polymerized outward toward the 5' terminus. The latter possibility seemed unlikely because: 1) the high degree of cooperativity in RecA protein binding under these conditions makes it unlikely that polymerization would randomly stop in the middle of the ssDNA tails, and 2) tracts of SSB protein were found at the ends of numerous tails following extended incubations, indicating that these molecules were not intermediates in the reaction. However, to exclude this possibility, it was necessary to localize the RecA protein nucleation sites along the ssDNA tails.

If RecA protein polymerized 5'→3' along the ssDNA tails, and if nucleation was occurring within the ssDNA tail, then examination of ssDNA tails in which RecA protein polymerization had not yet been completed should show a tract of RecA protein sandwiched between two SSB protein-covered tracts. However, if the polarity was 3'→5' and the nucleation was occurring at the ssDNA-dsDNA junction, no such molecules should be found. Following incubations of 2–4 min at

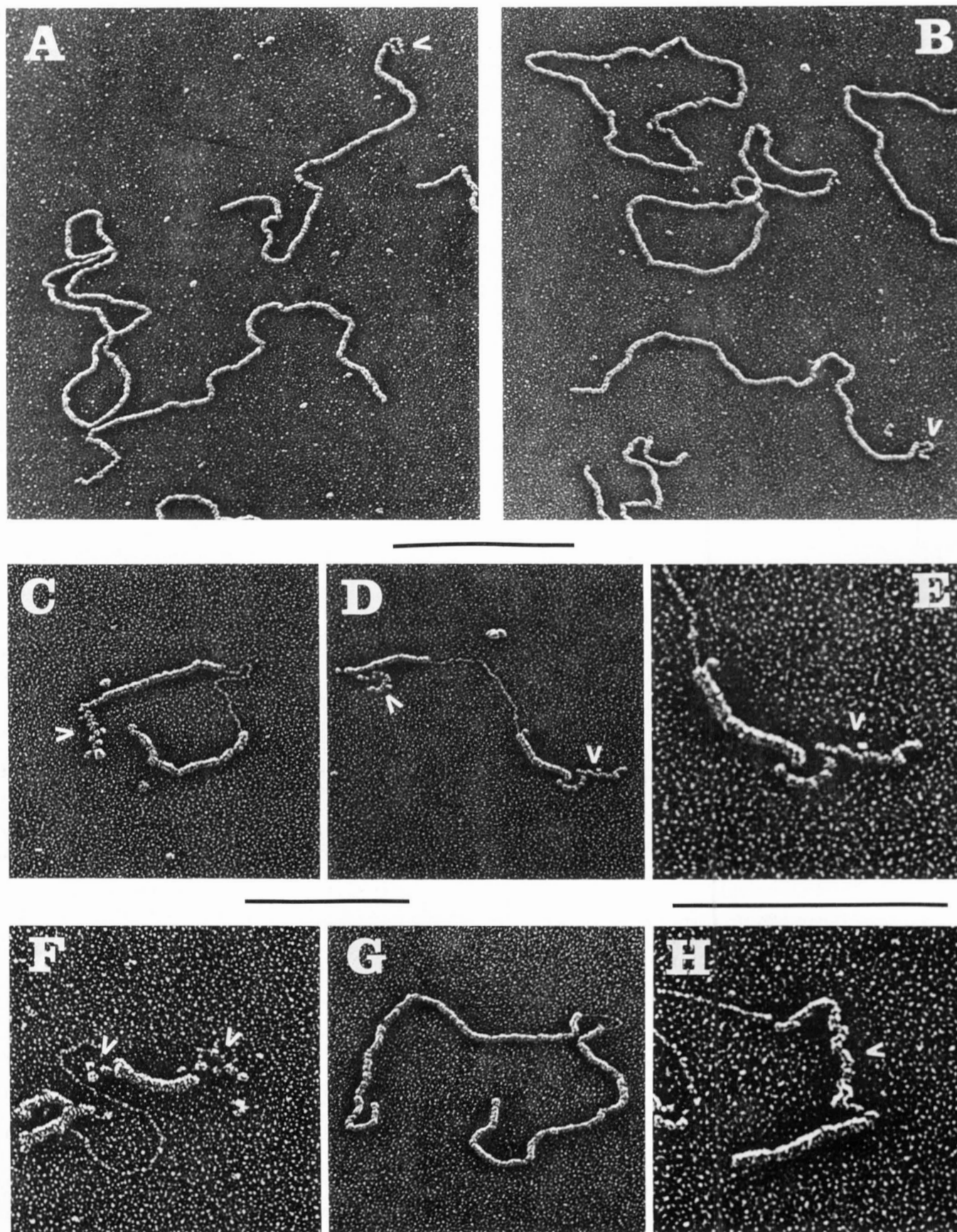


FIG. 1. Visualization of the polar assembly of RecA protein onto ssDNA. RecA protein was assembled onto several ssDNA substrates in the presence of Mg^{2+} , ATP, and SSB protein as described under "Materials and

21 °C (to slow the rate of polymerization), complexes were seen in which RecA protein polymerization appeared to be incomplete, and numerous single-stranded tails containing a RecA protein tract sandwiched between tracts of SSB protein were visualized (Fig. 1F). These results show that nucleation began within the length of the ssDNA tail and proceeded 5' → 3'.

RecA Protein Assembly onto Single-stranded DNA Tails with 3' Termini—To confirm that SSB-mediated assembly of RecA protein onto ssDNA occurs 5' → 3', assembly onto linear M13 dsDNA treated with λ -exonuclease was examined. This DNA contained long single-stranded tails with 3' termini. Experimental conditions were identical to those used with the exonuclease III-treated DNA. In contrast to the DNA which had been digested by exonuclease III, only 10% of the ssDNA tails generated by λ -exonuclease remained fully covered by SSB protein, while 58% were fully covered by RecA protein and 32% contained protein tracts typical of the two different types of protein coverage (Table I). In the last class, the tracts of RecA protein were consistently located at the end of the ssDNA tails (Fig. 1, G and H). Along with the data obtained from the experiments utilizing linear ssDNA and ssDNA tails with 5' ends, these results indicate that 3' ends are at least 10 times more likely to be covered by RecA protein as 5' ends under these conditions. In addition, these results show approximately a 2-fold preference for RecA protein nucleation at ssDNA-dsDNA junctions containing 5' termini over random initiation sites along the ssDNA tail.

When short incubations were carried out so that incompletely polymerized RecA protein tracts could be visualized, most (>50%) were found to begin at the ssDNA-dsDNA junction, but some were sandwiched between segments of ssDNA covered by SSB protein. To determine if the preferential nucleation seen at ssDNA-dsDNA junctions was due to the presence of a 5'-phosphate, RecA protein was assembled onto DNA which had been treated with λ -exonuclease and then dephosphorylated. No effect on RecA protein assembly was seen (data not shown).

SSB-mediated RecA Protein Assembly onto Gapped DNA—If RecA protein preferentially nucleates at ssDNA-dsDNA junctions containing 5' termini, then gaps in dsDNA should be rapidly covered over their entire length by RecA protein. Circular pUR222 dsDNA containing a single long gap in each molecule was used as the template for these experiments. Following incubation for 20 min at 37 °C, 83% of the molecules examined contained gaps that were completely covered by RecA protein, while 17% of the gaps were either completely covered by SSB protein or contained tracts typical of both types of protein coverage (Table I). It is worth noting that these results indicated an even greater preference for RecA protein nucleation at ssDNA-dsDNA junctions with 5' termini (almost 10-fold) than did those obtained with the λ -exonuclease-treated DNA.

DISCUSSION

We have shown that, when RecA protein binds to ssDNA in the presence of ATP, Mg^{2+} , and SSB protein to form

initiation complexes, it assembles in the 5' → 3' direction. In addition, we observed a strong preference for nucleation of RecA protein assembly at ssDNA-dsDNA junctions containing 5' termini.

The finding of polarity in the earliest phase of strand exchange correlates very well with the polarity seen in branch migration. Although numerous DNA substrates have been used to examine the strand exchange reaction, the most common are circular (+)-ssDNA circles and linear dsDNA from bacteriophages. In these model systems, the ssDNA invades the dsDNA 5' → 3' relative to the initiation complex (5–7), the same direction in which we have shown that RecA protein assembles onto ssDNA. The “treadmilling” model of branch migration (8, 30) fits well with our results. In this model, it is proposed that RecA protein assembles into filaments along ssDNA by a head-to-tail polymerization at one end of the filament. RecA protein concomitantly depolymerizes from the other end of the filament. The RecA protein-covered segment of ssDNA is the region which is actively involved in strand exchange. The overall movement of RecA protein along the ssDNA due to the treadmilling drives branch migration in the same direction. Since strand exchange is known to occur 5' → 3' with respect to the initiation complex (5–7), this model presupposes a 5' → 3' polarity of RecA protein assembly. Although an ATP-regenerating system was used here to avoid disassembly of the filaments, it is known that following sufficient ATP hydrolysis, RecA protein is released from these complexes (31). Further study is required to understand the molecular details of this disassembly.

Although the assembly of RecA protein onto ssDNA appears to be an important step in both homologous recombination (8, 12) and induction of the SOS response (32), very little is known about the relationship between the templates used for *in vitro* studies and those involved in the cell. It has been proposed that the DNA involved in homologous recombination *in vivo* probably more closely resembles a dsDNA circle (*i.e.* no readily available ends) and a gapped dsDNA than the templates more commonly used *in vitro* (33). The results presented here demonstrate that gaps are a very good substrate for the formation of initiation complexes *in vitro*, and this may explain why gaps readily promote recombination *in vivo* (34, 35). In addition, there is evidence that gaps may be important in the induction of the SOS response (for review, see Ref. 36).

In addition to exonucleolytic extension of nicks into gaps, ssDNA can be generated in the cell in several other ways: exonucleases can act on severed dsDNA to generate ssDNA tails with either 5' or 3' ends, helicase II (the *uvrD* product) and the rep protein unwind dsDNA, displacing 5' and 3' ends, respectively, and *E. coli* DNA polymerase I can displace ssDNA tails with 5' ends (for reviews, see Refs. 37 and 38). Meselson and Radding (39) have suggested that strand displacement by DNA polymerase I may play an important role in initiating recombination. Although our results suggest that a displaced 3' ssDNA end will be covered by RecA protein much faster than a displaced 5' end, strand exchange involving a displaced ssDNA tail with a 3' end would rapidly run

Methods” and in the text. When RecA protein was assembled onto M13mp7 ssDNA (3:1 ratio of linear to circular molecules) (A and B), it rapidly displaced all of the SSB protein from the circular DNA, whereas the linear filaments almost always contained a tract of SSB protein at one end (*carats*), indicating that the assembly was unidirectional. To determine the polarity of this assembly, linear M13 dsDNAs containing long ssDNA tails with 5' ends (C–F) or 3' ends (G and H) were used as templates. Under these conditions, RecA protein did not bind to the dsDNA which appears as a thin fiber connecting the protein filaments. Carats indicate SSB protein-covered tracts. Micrographs were taken at a 45° tilt to increase contrast; therefore, the magnifications are only approximate. Bars equal 0.5 μ m (A and B) and 0.25 μ m (C–H).

off the end of the tail due to the polarity of the strand exchange reaction. Despite the key role which ssDNA-RecA protein filaments play in both recombination and induction of the SOS response, little attention has been given to the effects that different ssDNA templates may have on these reactions. As this work shows, it will be important to address this relationship in order to completely understand the activities of the RecA protein *in vivo*.

REFERENCES

1. Radding, C. M. (1983) *Annu. Rev. Genet.* **16**, 405-437
2. Cox, M. M., and Lehman, I. R. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 3433-3437
3. DasGupta, C., Shibata, T., Cunningham, R. P., and Radding, C. M. (1980) *Cell* **22**, 437-446
4. Cunningham, R. P., Wu, A. M., Shibata, T., DasGupta, C., and Radding, C. M. (1981) *Cell* **24**, 213-223
5. Cox, M. M., and Lehman, I. R. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 6018-6022
6. West, S. C., Cassuto, E., and Howard-Flanders, P. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 6149-6153
7. Kahn, R., Cunningham, R. P., DasGupta, C., and Radding, C. M. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 4786-4790
8. Cox, M. M., Morrical, S. W., and Neuendorf, S. K. (1984) *Cold Spring Harbor Symp. Quant. Biol.* **49**, 525-533
9. Dunn, K., Chrysogelos, S., and Griffith, J. (1982) *Cell* **28**, 757-765
10. Flory, J., and Radding, C. M. (1982) *Cell* **28**, 747-756
11. Koller, Th., DiCapua, E., and Stasiak, A. (1983) in *Mechanisms of DNA Replication and Recombination* (Cozzarelli, N. R., ed) pp. 723-729, Alan R. Liss, Inc. New York
12. Radding, C. M., Flory, J., Wu, A. M., Kahn, R., DasGupta, C., Gonda, D., Bianchi, M., and Tsang, S. S. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 821-828
13. Kahn, R., and Radding, C. M. (1984) *J. Biol. Chem.* **259**, 7495-7509
14. Flory, J., Tsang, S. S., and Radding, C. M. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81**, 7026-7030
15. Bianchi, M., DasGupta, C., and Radding, C. M. (1983) *Cell* **34**, 931-939
16. Gonda, D. K., and Radding, C. M. (1983) *Cell* **34**, 647-654
17. Riddles, P. W., and Lehman, I. R. (1985) *J. Biol. Chem.* **260**, 165-169
18. Riddles, P. W., and Lehman, I. R. (1985) *J. Biol. Chem.* **260**, 170-173
19. West, S. C., Cassuto, E., and Howard-Flanders, P. (1982) *Mol. Gen. Genet.* **186**, 333-338
20. Cox, M. M., Soltis, D. A., Livneh, Z., and Lehman, I. R. (1983) *J. Biol. Chem.* **258**, 2577-2585
21. Griffith, J. D., Harris, L. D., and Register, J. C. (1984) *Cold Spring Harbor Symp. Quant. Biol.* **49**, 553-559
22. Chrysogelos, S., and Griffith, J. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79**, 5803-5807
23. Griffith, J., and Shores, C. G. (1985) *Biochemistry* **24**, 158-162
24. Chase, J. W., Whittier, R. F., Auerbach, J., Sancar, A., and Rupp, W. D. (1980) *Nucleic Acids Res.* **8**, 3215-3227
25. Rogers, S. G., and Weiss, B. (1980) *Methods Enzymol.* **65**, 201-211
26. Modrich, P., and Zabel, D. (1976) *J. Biol. Chem.* **251**, 5866-5874
27. Greenfield, L., Simpson, L., and Kaplan, D. (1975) *Biochim. Biophys. Acta* **407**, 365-375
28. Chaconas, G., and Van de Sande, J. H. (1980) *Methods Enzymol.* **65**, 75-85
29. Griffith, J. D., and Christiansen, G. (1978) *Annu. Rev. Biophys. Bioeng.* **7**, 19-37
30. Howard-Flanders, P., West, S. C., and Stasiak, A. (1984) *Nature* **304**, 209-216
31. Cox, M. M., Soltis, D. A., Lehman, I. R., DeBrosse, C., and Benkovic, S. J. (1983) *J. Biol. Chem.* **258**, 2586-2592
32. Phizicky, E. M., and Roberts, J. W. (1981) *Cell* **25**, 259-267
33. West, S. C., Cassuto, E., and Howard-Flanders, P. (1982) *Mol. Gen. Genet.* **187**, 209-217
34. Rupp, W. D., Wilde, C. E., Reno, D. L., and Howard-Flanders, P. (1971) *J. Mol. Biol.* **61**, 25-44
35. Ganesan, A. K. (1974) *J. Mol. Biol.* **87**, 103-119
36. Little, J. W., and Mount, D. W. (1982) *Cell* **29**, 11-22
37. Kornberg, A. (1980) *DNA Replication* W. H. Freeman and Co., San Francisco
38. Taucher-Scholz, G., Abdel-Monem, M., and Hoffmann-Berling, H. (1983) in *Mechanisms of DNA Replication and Recombination* (Cozzarelli, N. R., ed.) pp. 65-76, Alan R. Liss, Inc. New York
39. Meselson, M. S., and Radding, C. M. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72**, 358-361